A method of encoding input signals (I, r) to generate encoded data (100) is provided. The method involves processing the input signals (I, r) to determine first parameters (φ₁, φ₂) describing relative phase difference and temporal difference between the signals (I, r), and applying these first parameters (φ₁, φ₂) to process the input signals to generate intermediate signals. The method involves processing the intermediate signals to determine second parameters (α; IID, p) describing angular rotation of the first intermediate signals to generate a dominant signal (m) and a residual signal (s), the dominant signal (m) having a magnitude or energy greater than that of the residual signal (s). These second parameters are applicable to process the intermediate signals to generate the dominant (m) and residual (s) signals. The method also involves quantizing the first parameters, the second parameters, and dominant and residual signals (m, s) to generate corresponding quantized data for subsequent multiplexing to generate the encoded data (100).
**U.S. PATENT DOCUMENTS**

<table>
<thead>
<tr>
<th>Patent Number</th>
<th>Date</th>
<th>Inventors</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,636,324 A</td>
<td>6/1997</td>
<td>Teh et al.</td>
<td>704/226</td>
</tr>
<tr>
<td>5,682,461 A</td>
<td>10/1997</td>
<td>Silke et al.</td>
<td>704/205</td>
</tr>
<tr>
<td>7,181,019 B2</td>
<td>2/2007</td>
<td>Booibaert et al.</td>
<td>381/23</td>
</tr>
<tr>
<td>7,272,556 B1</td>
<td>9/2007</td>
<td>Aguilar et al.</td>
<td>704/230</td>
</tr>
</tbody>
</table>

**OTHER PUBLICATIONS**


* cited by examiner
FIG. 9
The present invention relates to methods of coding data, for example to a method of coding audio and/or image data utilizing variable angle rotation of data components. Moreover, the invention also relates to decoders employing such methods, and to decoders operable to decode data generated by these encoders. Furthermore, the invention is concerned with encoded data communicated via data carriers and/or communication networks, the encoded data being generated according to the methods.

Numerous contemporary methods are known for encoding audio and/or image data to generate corresponding encoded output data. An example of a contemporary method of encoding audio is MPEG-1 Layer III known as MP3 and described in ISO/IEC JTC1/SC29/WG11 MPEG, IS 11172-3, Information Technology—Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s, Part 3: Audio, MPEG-1, 1992. Some of these contemporary methods are arranged to improve coding efficiency, namely provide enhanced data compression, by employing mid-side (M/S) stereo coding or sum/difference stereo coding as described by J. D. Johnston and A. J. Ferreira, "Sum-difference stereo transform coding", in Proc. IEEE, Int. Conf. Acoust., Speech and Signal Proc., San Francisco, Calif., March 1992, pp. 569-572.

In M/S coding, a stereo signal comprises left and right signals $l[n]$, $r[n]$ respectively which are coded as a sum signal $m[n]$ and a difference signal $s[n]$, for example by applying processing as described by Equations 1 and 2 (Eq. 1 and 2):

$$m[n] = l[n] + r[n]$$
$$s[n] = l[n] - r[n]$$

When the signals $l[n]$ and $r[n]$ are almost identical, the M/S coding is capable of providing significant data compression on account of the difference signal $s[n]$ approaching zero and thereby conveying relatively little information whereas the sum signal effectively includes most of the signal information content. In such a situation, a bit rate required to represent the sum and difference signals is close to half that required for independently coding the signals $l[n]$ and $r[n]$.

Equations 1 and 2 are susceptible to being represented by way of a rotation matrix as in Equation 3 (Eq. 3):

$$\begin{pmatrix}
m[n] \\
s[n]
\end{pmatrix} =
\begin{pmatrix}
\cos(\alpha) & \sin(\alpha) \\
-\sin(\alpha) & \cos(\alpha)
\end{pmatrix}
\begin{pmatrix}
l[n] \\
r[n]
\end{pmatrix}$$

wherein $\alpha$ is a constant scaling coefficient often used to prevent clipping.

Whereas Equation 3 effectively corresponds to a rotation of the signals $l[n]$, $r[n]$ by an angle of $45^\circ$, other rotation angles are possible as provided in Equation 4 (Eq. 4) wherein $\alpha$ is a rotation angle applied to the signals $l[n]$, $r[n]$ to generate corresponding coded signals $m'[n]$, $s'[n]$ hereinafter described as relating to dominant and residual signals respectively:

$$\begin{pmatrix}
m'[n] \\
s'[n]
\end{pmatrix} =
\begin{pmatrix}
\cos(\alpha) & \sin(\alpha) \\
-\sin(\alpha) & \cos(\alpha)
\end{pmatrix}
\begin{pmatrix}
l[n] \\
r[n]
\end{pmatrix}$$

The angle $\alpha$ is beneficially made variable to provide enhanced compression for a wide class of signals $l[n]$, $r[n]$ by reducing information content present in the residual signal $s[n]$ and concentrating information content in the dominant signal $m'[n]$, namely minimize power in the residual signal $s'[n]$ and consequently maximize power in the dominant signal $m'[n]$.

Coding techniques represented by Equations 1 to 4 are conventionally not applied to broadband signals but to sub-signals each representing only a smaller part of a full bandwidth used to convey audio signals. Moreover, the techniques of Equations 1 to 4 are also conventionally applied to frequency domain representations of the signals $l[n]$, $r[n]$.

In a published U.S. Pat. No. 5,621,855, there is described a method of sub-band coding a digital signal having first and second signal components, the digital signal being sub-band coded to produce a first sub-band signal having a first q-sample signal block in response to the first signal component, and a second sub-band signal having a second q-sample signal block in response to the second signal component, the first and second sub-band signals being in the same sub-band and the first and second signal blocks being time equivalent.

The first and second signal blocks are processed to obtain a minimum distance value between point representations of time-equivalent samples. When the minimum distance value is less than or equal to a threshold distance value, a composite block composed of q samples is obtained by adding the respective pairs of time-equivalent samples in the first and second signal blocks together after multiplying each of the samples of the first block by $\cos(\alpha)$ and each of the samples of the second signal block by $-\sin(\alpha)$.

Although application of the aforementioned rotation angle $\alpha$ is susceptible to eliminating many disadvantages of M/S coding where only a 45° rotation is employed, such approaches are found to be problematic when applied to groups of signals, for example stereo signal pairs, when considerable relative mutual phase or time offsets in these signals occur. The present invention is directed at addressing this problem.

An object of the present invention is to provide a method of encoding data.

According to a first aspect of the present invention, there is provided a method of encoding a plurality of input signals $l$, $r$ to generate corresponding encoded data, the method comprising steps of:

(a) processing the input signals $l$, $r$ to determine first parameters ($\phi_2$) describing at least one of relative phase difference and temporal difference between the signals $l$, $r$, and applying these first parameters ($\phi_2$) to process the input signals to generate corresponding intermediate signals;

(b) processing the intermediate signals and/or the input signals $l$, $r$ to determine second parameters describing rotation of the intermediate signals required to generate a dominant signal (m) and a residual signal (s), said dominant signal (m) having a magnitude or energy greater than that of the residual signal (s), and applying these second parameters to process the intermediate signals to generate the dominant (m) and residual (s) signals;
(c) quantizing the first parameters, the second parameters, and
encoding at least a part of the dominant signal (m) and
the residual signal (s) to generate corresponding quantized
data; and
(d) multiplexing the quantized data to generate the encoded
data.

The invention is of advantage in that it is capable of providing
for more efficient encoding of data.

Preferably, in the method, only a part of the residual
signal (s) is included in the encoded data. Such partial inclu-
sion of the residual signal (s) is capable of enhancing data
compression achievable in the encoded data.

More preferably, in the method, the encoded data also
includes one or more parameters indicative of parts of the
residual signal included in the encoded data. Such indicative
parameters are susceptible to rendering subsequent decoding
of the encoded data less complex.

Preferably, steps (a) and (b) of the method are implemented
by complex rotation with the input signals (lln), (lkn),
represented in the frequency domain (f[kn], f[k]). Implementation
of complex rotation is capable of more efficiently coping with
relative temporal and/or phase differences arising between
the plurality of input signals. More preferably, steps (a) and
(b) are performed in the frequency domain or a sub-bands
domain. “Sub-band” is to be construed to be a frequency
region smaller than a full frequency bandwidth required for a
signal.

Preferably, the method is applied in a sub-part of a full
frequency range encompassing the input signals (l, r).
More preferably, other sub-parts of the full frequency range are
encoded using alternative encoding techniques, for example
conventional MS encoding as described in the foregoing.

Preferably, the method includes an additional step after
step (e) of losslessly coding the quantized data to provide the
data for multiplexing in step (d) to generate the encoded data.
More preferably, the lossless coding is implemented using
Huffman coding. Utilizing lossless coding enables poten-
tially higher audio quality to be achieved.

Preferably, the method includes a step of manipulating the
residual signal (s) by discarding perceptually non-relevant
time-frequency information present in the residual signal (s),
said manipulated residual signal (s) contributing to the
encoded data (100), and said perceptually non-relevant infor-
mation corresponding to selected portions of a spectro-
temporal representation of the input signals. Discarding percep-
tually non-relevant information enables the method to
provide a greater degree of data compression in the encoded
data.

Preferably, in step (b) of the method, the second parameters
(α; IID, p) are derived by minimizing the magnitude or
energy of the residual signal (s). Such an approach is com-
putationally efficient for generating the second parameters in
comparison to alternative approaches to deriving the param-
eters.

Preferably, in the method, the second parameters (α; IID,
p) are represented by way of inter-channel intensity differ-
ence parameters and coherence parameters (II, p). Such
implementation of the method is capable of providing back-
ward compatibility with existing parametric stereo encoding
and associated decoding hardware or software.

Preferably, in steps (c) and (d) of the method, the encoded
data is arranged in layers of significance, said layers including
a base layer conveying the dominant signal (m), a first
enhancement layer including first and/or second parameters
corresponding to stereo imparting parameters, a second
enhancement layer conveying a representation of the residual
signal (s). More preferably, the second enhancement layer is
further subdivided into a first sub-layer for conveying most
relevant time-frequency information of the residual signal (s)
and a second sub-layer for conveying less relevant time-
frequency information of the residual signal (s). Representa-
tion of the input signals by these layers, and sub-layers as
described here, is capable of enhancing robustness to transmission
errors of the encoded data and rendering it backward compatible
with simpler decoding hardware.

According to a second aspect of the invention, there is
provided an encoder for encoding a plurality of input signals
(l, r) to generate corresponding encoded data, the encoder
comprising:
(a) first processing means for processing the input signals (l,
r) to determine first parameters (φ1) describing at least one
of relative phase difference and temporal difference
between the signals (l, r), the first processing means being
operable to apply these first parameters (φ1) to process the
input signals to generate corresponding intermediate
signals;
(b) second processing means for processing the intermediate
signals to determine second parameters describing rotation
of the intermediate signals required to generate a dominant
signal (m) and a residual signal (s), said dominant signal
(m) having a magnitude or energy greater than that of the
residual signal (s), the second processing means being
operable to apply these second parameters to process the
intermediate signals to generate at least the dominant (m)
and residual (s) signals;
(c) quantizing means for quantizing the first parameters (φ2),
the second parameters (α; IID, p), and at least a part of the
dominant signal (m) and the residual signal (s) to generate
corresponding quantized data; and
(d) multiplexing means for multiplexing the quantized data to
generate the encoded data.

The encoder is of advantage in that it is capable of providing
for more efficient encoding of data.

Preferably, the encoder comprises processing means for
manipulating the residual signal (s) by discarding perceptually
non-relevant time-frequency information present in the
residual signal (s), said transformed residual signal (s) con-
tributing to the encoded data (100) and said perceptually non-relevant information corresponding to selected portions of a spectro-
temporal representation of the input signals. Discard-
ing perceptually non-relevant information enables the encoder to provide a greater degree of data compression in the encoded
data.

According to a third aspect of the present invention, there is
provided a method of decoding encoded data to regenerate
corresponding representations of a plurality of input signals
(l, r), said input signals (l, r) being previously encoded to
generate said encoded data, the method comprising steps of:
(a) demultiplexing the encoded data to generate correspond-
ing quantized data;
(b) processing the quantized data to generate corresponding
first parameters (φ2), second parameters, and at least a
dominant signal (m) and a residual signal (s), said domi-
nant signal (m) having a magnitude or energy greater than
that of the residual signal (s);
(c) rotating the dominant (m) and residual (s) signals by
applying the second parameters to generate corresponding
intermediate signals; and
(d) processing the intermediate signals by applying the first
parameters (φ2) to regenerate said representations of said
input signals (l, r), the first parameters (φ1) describing at
least one of relative phase difference and temporal difference
between the signals (l, r).
The method provides an advantage of being capable of efficiently decoding data which has been efficiently coding using a method according to the first aspect of the invention.

Preferably, step (b) of the method includes a further step of appropriately supplementing missing time-frequency information of the residual signal (s) with a synthetic residual signal derived from the dominant signal (m). Generation of the synthetic signal is capable of resulting in efficient decoding of encoded data.

Preferably, in the method, the encoded data includes parameters indicative of which parts of the residual signal (s) are encoded into the encoded data. Inclusion of such indicative parameters is capable of rendering decoding for efficient and less computationally demanding.

According to a fourth aspect of the present invention, there is provided a decoder for decoding encoded data to regenerate corresponding representations of a plurality of input signals (m, s), said input signals (l, r) being previously encoded to generate the encoded data, the decoder comprising:
(a) de-multiplexing means for de-multiplexing the encoded data to generate corresponding quantized data;
(b) first processing means for processing the quantized data to generate corresponding first parameters (\( \phi_1 \)), second parameters, and at least a dominant signal (m) and a residual signal (s), said dominant signal (m) having a magnitude or energy greater than that of that residual signal (s);
(c) second processing means for rotating the dominant (m) and residual (s) signals by applying the second parameters to generate corresponding intermediate signals; and
(d) third processing means for processing the intermediate signals by applying the first parameters (\( \phi_1 \)) to regenerate said representations of the input signals (l, r), the first parameters (\( \phi_1 \)) describing at least one of relative phase difference and temporal difference between the signals (l, r).

Preferably, the second processing means is operable to generate a supplementary synthetic signal derived from the decoded dominant signal (m) for providing information missing from the decoded residual signal.

According to a fifth aspect of the invention, there is provided encoded data generated according to the method of the first aspect of the invention, the data being at least one of recorded on a data carrier and communicable via a communication network.

According to a sixth aspect of the invention, there is provided software for executing the method of the first aspect of the invention on computing hardware.

According to a seventh aspect of the invention, there is provided software for executing the method of the third aspect of the invention on computing hardware.

According to an eighth aspect of the invention, there is provided encoded data at least one of recorded on a data carrier and communicable via a communication network, said data comprising a multiplex of quantizing first parameters, quantized second parameters, and quantized data corresponding to at least a part of a dominant signal (m) and a residual signal (s), wherein the dominant signal (m) has a magnitude or energy greater than the residual signal (s), said dominant signal (m) and said residual signal (s) being derivable by rotating intermediate signals according to the second parameters, said intermediate signals being generated by processing a plurality of input signals to compensate for relative phase and/or temporal delays therebetween as described by the first parameters.

It will be appreciated that features of the invention are susceptible to being combined in any combination without departing from the scope of the invention as defined in the accompanying claims.

Embodiments of the invention will now be described, by way of example only, with reference to the following diagrams wherein:

FIG. 1 is an illustration of sample sequences for signals l[n], r[n] subject to relative mutual time and phase delays;

FIG. 2 is an illustration of application of a conventional MS transform pursuant to Equations 1 and 2 applied to the signals of FIG. 1 to generate corresponding sum and difference signals m[n], s[n];

FIG. 3 is an illustration of application of a rotation transform pursuant to Equation 4 applied to the signals of FIG. 1 to generate corresponding dominant m[n] and residual s[n] signals;

FIG. 4 is an illustration of application of a complex rotation transform according to the invention pursuant to Equations 5 to 15 to generate corresponding dominant m[n] and residual s[n] signals wherein the residual signal is of relatively small amplitude despite the signals of FIG. 1 having relative mutual phase and time delay;

FIG. 5 is a schematic diagram of a encoder according to the invention;

FIG. 6 is a schematic diagram of a decoder according to the invention, the encoder being compatible with the encoder of FIG. 5;

FIG. 7 is a schematic diagram of a parametric stereo decoder;

FIG. 8 is a schematic diagram of an enhanced parametric stereo encoder according to the invention; and

FIG. 9 is a schematic diagram of an enhanced parametric stereo decoder according to the invention, the decoder being compatible with the encoder of FIG. 9.

In overview, the present invention is concerned with a method of coding data which represents an advance to MS coding methods described in the foregoing employing a variable rotation angle. The method is devised by the inventors to be better capable of coding data corresponding to groups of signals subject to considerable phase and/or time offset. Moreover, the method provides advantages in comparison to conventional coding techniques by employing values of the rotation angle \( \alpha \) which can be used when the signals l[n], r[n] are represented by their equivalent complex-valued frequency domain representations \( l[k] \), \( r[k] \) respectively.

The angle \( \alpha \) can be arranged to be real-valued and a real-valued phase rotation applied to mutually "cohere" the l[n], r[n] signals to accommodate mutual temporal and/or phase delays between these signals. However, use of complex values for the rotation angle \( \alpha \) renders the present invention easier to implement. Such an alternative approach to implementing rotation by angle \( \alpha \) is to be construed to be within the scope of the present invention.

Frequency-domain representations of the aforesaid time-domain signals l[n], r[n] are preferably derived by applying a method of encoding data comprising a multiplex of quantizing first parameters, quantized second parameters, and quantized data corresponding to at least a part of a dominant signal (m) and a residual signal (s), wherein the dominant signal (m) has a magnitude or energy greater than the residual signal (s), said dominant signal (m) and said residual signal (s) being derivable by rotating intermediate signals according to the second parameters, said intermediate signals being generated by processing a plurality of input signals to compensate for relative phase and/or temporal delays therebetween as described by the first parameters:

\[
l[n] \rightarrow \{l[n+\alpha]\} h[n]\]
\[
r[n] \rightarrow \{r[n+\alpha]\} h[n]\]
\]

wherein

q = a frame index such that q = 0, 1, 2, . . . to indicate consecutive signal frames;
H—a hop-size or update-size; and
n—a time index having a value in a range of 0 to L-1, where
a parameter L is equivalent to the length of a window h[n].
The windowed signals l[n], r[n] are transformable to the
frequency domain by using a Discrete Fourier Transform (DFT),
or functionally equivalent transform, as described in
Equations 7 and 8 (Eq. 7 and 8):

\[ l[k] = \sum_{n=0}^{N-1} l[n] \exp\left(\frac{-2\pi kn}{N}\right) \]
\[ r[k] = \sum_{n=0}^{N-1} r[n] \exp\left(\frac{-2\pi kn}{N}\right) \]

wherein a parameter N represents a DFT length such that
N≤L. On account of the DFT of a real-valued sequence being
symmetrical, only the first N/2+1 points are preserved after
the transform. In order to preserve signal energy when implement-
ing the DFT, the following scaling as described in Equations
9 and 10 (Eq. 9 and 10) is preferably employed:

\[ l[0] = \frac{l[0]}{2} \]
\[ r[0] = \frac{r[0]}{2} \]

The method of the invention performs signal processing
operations as described by Equation 11 (Eq. 11) to convert the
frequency domain signal representations l[k], r[k] in Equa-
tions 7 and 8 to corresponding rotated sum and difference
signals m*[k], s*[k] in the frequency domain:

\[
\begin{pmatrix}
  m^*[k] \\ s^*[k]
\end{pmatrix} =
\begin{pmatrix}
  \cos(\alpha) & \sin(\alpha) \\
  -\sin(\alpha) & \cos(\alpha)
\end{pmatrix}
\begin{pmatrix}
  e^{j\phi_1} & 0 \\
  0 & e^{j\phi_2}
\end{pmatrix}
\begin{pmatrix}
  l[k] \\ r[k]
\end{pmatrix}
\]

wherein
\[ \alpha \]—real-valued variable rotation angle;
\[ \phi_1 \]—a common angle used to maximise the continuation of
signals over associated boundaries; and
\[ \phi_2 \]—an angle used to minimize the energy of the residual
signal s*[k] by phase-rotating the right signal r[k].

Use of the angle \( \phi_1 \) is optional. Moreover, rotations pursuant
to Equation 11 are preferably executed on a frame-by-
frame basis, namely dynamically in frame steps. However,
such dynamic changes in rotation from frame-to-frame can
potentially cause signal discontinuities in the sum signal
m*[k] which can be at least partially removed by suitable
selection of the angle \( \phi_1 \).

Furthermore, the frequency range \( k=0 \ldots N/2+1 \) of Equation
11 is preferably divided into sub-ranges, namely regions.
For each region during encoding, its corresponding angle
parameters \( \alpha, \phi_1 \) and \( \phi_2 \) are then independently determined,
coded and then transmitted or otherwise conveyed to a
codec for subsequent decoding. By arranging for the
frequency range to be sub-divided, signal properties can be
better captured during encoding resulting potentially in
higher compression ratios.

After implementing mappings pursuant to Equations 7 to
11, the signals m*[k], s*[k] are subjected to an inverse Dis-
crete Fourier Transform as described in Equations 12 and 13
(Eq. 12 & 13):

\[ m[n] = \sum_{k=0}^{N-1} m[k] \exp\left(\frac{j2\pi kn}{N}\right) \]
\[ s[n] = \sum_{k=0}^{N-1} s[k] \exp\left(\frac{j2\pi kn}{N}\right) \]

wherein
\[ m[n] \]—dominant time-domain representation; and
\[ s[n] \]—residual (difference) time-domain representation.

The dominant and residual representations are then con-
verted in the method to representations on a windowed basis
to which overlap is applied as provided by processing opera-
tions as described by Equations 14 and 15 (Eq. 14 and 15):

\[
\begin{array}{l}
  \text{m}[n]=m[n]+q[H]*2\text{Re}\{m[n]*h[n]\} \\
  \text{s}[n]=s[n]+q[H]*2\text{Re}\{s[n]*h[n]\}
\end{array}
\]

Alternatively, processing operations of the method of the invention as described by Equations 5 to 15 are susceptible, at
least in part, to being implemented in practice by employing
complex-modulated filter banks. Digital processing applied
in computer processing hardware can be employed to imple-
ment the invention.

In order to illustrate the method of the invention, a signal
processing example of the invention will now be described.
For the example, two temporal signals are used as initial
signals to be processed using the method, the two signals
being defined by Equations 16 and 17 (Eq. 16 and 17):

\[
\begin{align*}
  l[n] &= 0.5 \cos(0.32n+0.4)+0.05z_1[n]+0.06z_2[n] \\
  r[n] &= 0.25 \cos(0.32n+1.8)+0.03z_1[n]+0.05z_2[n]
\end{align*}
\]

wherein \( z_1[n], z_2[n] \) and \( z_3[n] \) are mutually independent white
noise sequences of unity variance. In order to better appreciate
operation of the method of the invention, portions of the
signals \( l[n], r[n] \) described by Equations 16 and 17 are shown
in FIG. 1.

In FIG. 2, M/S transform signals \( m[n] \) and \( s[n] \) are illus-
trated, these transform signals being derived from the signals
\( l[n], r[n] \) of Equations 16 and 17 by conventional processing
pursuant to Equations 1 and 2. It will be seen from FIG. 2 that
such a conventional approach to generating the signals \( m[n] \)
and \( s[n] \) from the signals of Equations 16 and 17 results in the
energy of the residual signal \( s[n] \) being higher than the energy
of the input signal \( r[n] \) in Equation 17. Clearly, conventional
M/S transform signal processing applied to the signals of
Equations 16 and 17 is ineffective at resulting in signal com-
pression because the signal \( s[n] \) is not of negligible magni-
itude.

By employing a rotation transform as described by Equa-
tion 4, it is possible for the example signals \( l[n], r[n] \) to reduce
the residual energy in their corresponding residual signal \( s[n] \)
and correspondingly enhance their dominant signal \( m[n] \) as
illustrated in FIG. 3. Although the rotation approach of Equa-
tion 4 is capable of performing better than conventional M/S
processing as presented in FIG. 2, it is found by the inventors
to be unsatisfactory when the signals \( l[n], r[n] \) are subject to
relative mutual phase and/or time shifts.
When the sample signals \( l[n], r[n] \) of Equations 16 and 17 are subjected to transformation to the frequency domain, then subjected to a complex optimizing rotation pursuant to the Equations 5 to 15, it is feasible to reduce the energy of the residual signal \( s[n] \) to a comparatively small magnitude as illustrated in FIG. 4.

Embeddings of encoder hardware operable to implement signals processing as described by Equations 5 to 15 will next be described.

In FIG. 5, there is shown an encoder according to the invention indicated generally by 10. The encoder 10 is operable to receive left (l) and right (r) complementary input signals and encode these signals to generate an encoded bit-stream (bs) 100. Moreover, the encoder 10 includes a phase rotation unit 20, a signal rotation unit 30, a time/frequency selector 40, a first coder 50, a second coder 60, a parameter quantizing processing unit (Q) 70 and a bit-stream multiplexer unit 80.

The input signals l, r are coupled to inputs of the phase rotation unit 20 whose corresponding outputs are connected to the signal rotation unit 30. Dominant and residual signals of the signal rotation unit 30 are denoted by m, s respectively.

The dominant signal m is conveyed via the first coder 50 to the multiplexer unit 80. Moreover, the residual signal \( s \) is coupled via the time/frequency selector 40 to the second coder 60 and thereafter to the multiplexer unit 80. Angle parameter outputs \( \phi_1, \phi_2 \) from the phase rotation unit 20, a time/frequency processing output \( \phi_3 \), and from the signal rotation unit 30 via the processing unit 70 to the multiplexer unit 80. Additionally, an angle parameter output \( \alpha \) is coupled from the signal rotation unit 30 via the processing unit 70 to the multiplexer unit 80. The multiplexer unit 80 comprises the aforementioned encoded bit stream output (bs) 100.

In operation, the phase rotation unit 20 applies processing to the signals l, r to compensate for relative phase differences therebetween and thereby generate the parameters \( \phi_1, \phi_2 \) wherein the parameter \( \phi_2 \) is representative of such relative phase difference, the parameters \( \phi_1, \phi_2 \) being passed to the processing unit 70 for quantizing and thereby including as corresponding parameter data in the encoded bit stream 100. The signals l, r compensated for relative phase difference pass to the signal rotation unit 30 which determines an optimized value for the angle \( \alpha \) to concentrate a maximum amount of signal energy in the dominant signal m and a minimum amount of signal energy in the residual signal s. The dominant and residual signals m, s then pass via the coders 50, 60 to be converted to a suitable format for inclusion in the bit stream 100. The processing unit 70 receives the angle signals \( \alpha, \phi_1, \phi_2 \) and multiplexes them together with the output from the coders 50, 60 to generate the bit-stream output (bs) 100. Thus, the bit-stream (bs) 100 thereby comprises a stream of data including representations of the dominant and residual signals m, s together with angle parameter data \( \alpha, \phi_1, \phi_2 \) wherein the parameter \( \phi_2 \) is essential and the parameters \( \phi_1 \) are optional but nevertheless beneficial to include.

The coders 50, 60 are preferably implemented as two mono audio encoders, or alternatively as one dual mono encoder. Optionally, certain parts of the residual signal s, for example identified when represented in a time-frequency plane, not perceptibly contributing to the bit stream 100 can be discarded in the time/frequency selector 40, thereby providing scalable data compression as elucidated in more detail below.

The encoder 10 is optionally capable of being used for processing the input signals (l, r) over a part of a full frequency range encompassing the input signals. Those parts of the input signals (l, r) not encoded by the encoder 10 are then in parallel encoded using other methods, for example using conventional M/S encoding as described in the foregoing. If required individual encoding of left (l) and right (r) input signals can be implemented if required.

The encoder 10 is susceptible to being implemented in hardware, for example as an application specific integrated circuit or group of such circuits. Alternatively, the encoder 10 can be implemented in software executing on computing hardware, for example on a proprietary software-driven signal processing integrated circuit or group of such circuits.

In FIG. 6, a decoder compatible with the encoder 10 is indicated generally by 200. The decoder 200 comprises a bit-stream demultiplexer 210, first and second decoders 220, 230, a processing unit 240 for de-quantizing parameters, a signal rotation decoder unit 250 and a phase rotation decoding unit 260 providing decoded outputs \( l', r' \) corresponding to the input signals l, r input to the encoder 10. The demultiplexer 210 is configured to receive the bit-stream (bs) 100 as generated by the encoder 10, for example conveyed from the encoder 10 to the decoder 200 by way of a data carrier, for example an optical disk data carrier such as a CD or DVD, and/or a communication network, for example the Internet. Demultiplexed outputs of the demultiplexer 210 are coupled to inputs of the decoders 220, 230 and to the processing unit 240. The first and second decoders 220, 230 comprise dominant and residual decoded outputs \( m', s' \) respectively which are coupled to the rotation decoder unit 250. Moreover, the processing unit 240 includes a rotation angle output \( \alpha' \) which is also coupled to the rotation decoder unit 250; the angle \( \alpha' \) corresponds to a decoded version of the aforementioned angle \( \alpha \) with regard to the encoder 10. Angle outputs \( \phi_{1}', \phi_{2}' \) correspond to decoded versions of the aforementioned angles \( \phi_{1}, \phi_{2} \) with regard to the encoder 10; these angle outputs \( \phi_{1}', \phi_{2}' \) are conveyed, together with decoded dominant and residual signal outputs from the rotation decoder unit 250 to the phase rotation decoding unit 260 which includes decoded outputs \( l', r' \) as illustrated.

In operation, the decoder 200 performs an inverse of encoding steps executed within the encoder 10. Thus, in the decoder 200, the bit-stream 100 is demultiplexed in the demultiplexer 210 to isolate data corresponding to the dominant and residual signals which are reconstituted by the decoders 220, 230 to generate the decoded dominant and residual signals \( m', s' \). These signals \( m', s' \) are then rotated according to the angle \( \alpha' \) and then corrected for relative phase using the angles \( \phi_{1}', \phi_{2}' \) to regenerate the left and right signals \( l', r' \). The angles \( \phi_{1}', \phi_{2}', \alpha' \) are regenerated from parameters demultiplexed in the demultiplexer 210 and isolated in the processing unit 240.

In the encoder 10, and hence also in the decoder 200, it is preferable to transmit in the bit-stream 100 an IID value and a coherence value \( \rho \) rather than the aforementioned angle \( \alpha \). The IID value is arranged to represent an inter-channel difference, namely denoting frequency and time variant magnitude differences between the left and right signals l, r. The coherence value \( \rho \) denotes frequency variant coherence, namely similarity, between the left and right signals l, r after phase synchronization. However, for example in the decoder 200, the angle \( \alpha \) is readily derivable from the IID and \( \rho \) values by applying Equation 18 (Eq. 18):

\[
\alpha = \frac{1}{2} \arctan \left( \frac{10^{m_p/10} \rho}{10^{m_p/10} - 1} \right)
\]

A parametric decoder is indicated generally by 400 in FIG. 7, this decoder 400 being complementary to the encoders
according to the present invention. The decoder 400 comprises a bit-stream demultiplexer 410, a decoder 420, a de-correlation unit 430, a scaling unit 440, a signal rotation unit 450, a phase rotation unit 460 and a de-quantizing unit 470. The demultiplexer unit 410 comprises an input for receiving the bit-stream signal (bs) 100 and four corresponding outputs for signals m, s, data, angle parameter data, IID data and coherence data ρ, these outputs are connected to the decoder 420 and to the de-quantizer unit 470 as shown. An output from the decoder 420 is coupled via the de-correlation unit 430 for regenerating a representation of the residual signal s’ for input to the scaling function 440. Moreover, a regenerated representation of the dominant signal m’ is conveyed from the decoder unit 420 to the scaling unit 440. The scaling unit 440 is also provided with IID and coherence data ρ from the de-quantizing unit 470. Outputs from the scaling unit 440 are coupled to the signal rotation unit 450 to generate intermediate output signals. These intermediate output signals are then corrected in the phase rotation unit 460 using the angles ϕ1, ϕ2, decoded in the de-quantizing unit 470 to regenerate representations of the left and right signals l, r.

The decoder 400 is distinguished from the decoder 200 of FIG. 6 in that the decoder 400 includes the decorrelation unit 430 for estimating the residual signal s’ based on the dominant signal m’ by the way of decorrelation processes executed within the de-correlation unit 430. Moreover, the amount of coherence between the left and right output signals l, r is determined by way of a scaling operation. The scaling operation is executed within the scaling unit 440 and is concerned with a ratio between the dominant signal m’ and the residual signal s’.

Referring next to FIG. 8, there is illustrated an enhanced encoder indicated generally by 500. The encoder 500 comprises a phase rotation unit 510 for receiving left and right input signals l, r respectively, a signal rotation unit 520, a time/frequency selector 530, and first and second coders 540, 550, 560 respectively. The quantizing unit 560 comprises a quantizer unit 560 and a multiplexer 570 including the bit-stream output (bs) 100. Angle outputs ϕ1, ϕ2 from the phase rotation unit 510 are coupled from the phase rotation unit 510 to the quantizing unit 560. Moreover, phase-corrected outputs from the phase rotation unit 510 are connected via the signal rotation unit 520 and the time/frequency selector 530 to generate dominant and residual signals m, s respectively, as well as IID and coherence parameter data. The IID and coherence parameter data are coupled to the quantizer unit 560 whereas the dominant and residual signals m, s are passed via the first and second coders 540, 550 to generate corresponding data for the multiplexer 570. The multiplexer 570 is also arranged to receive parameter data describing the angles ϕ1, ϕ2, the coherence parameter and the IID. The multiplexer 570 is operable to multiplex data from the coders 540, 550 and the quantizer unit 560 to generate the bit-stream (bs) 100.

In the encoder 500, the residual signal s is encoded directly into the bit-stream 100. Optionally, the time/frequency selector unit 530 is operable to determine which parts of the time/frequency plane of the residual signal s are encoded into the bit-stream (bs) 100, the unit 530 thereby determining a degree to which residual information is included in the bit-stream 100 and hence affecting a compromise between compression attainable in the encoder 500 and degree of information included within the bit-stream 100.

In FIG. 9, an enhanced parametric decoder is indicated generally by 600, the decoder 600 being complementary to the encoder 500 illustrated in FIG. 8. The decoder 600 comprises a demultiplexer unit 610, first and second decoders 620, 640 respectively, a de-correlation unit 630, a combiner unit 650, a scaling unit 660, a signal rotation unit 670, a phase rotation unit 680 and the de-quantizing unit 690. The demultiplexer unit 610 is coupled to receive the encoded bit-stream (bs) 100 and provide corresponding demultiplexed outputs to the first and second decoders 620, 640 and also to the de-multiplexer unit 690. The decoders 620, 640 in conjunction with the de-correlation unit 630 and the combiner unit 650 are operable to regenerate representations of the dominant and residual signals m, s respectively. These representations are subjected to scaling processes in the scaling unit 660 followed by rotations in the signal rotation unit 670 to generate intermediate signals which are then phase rotated in the rotation unit 680 in response to angle parameters generated by the de-quantizing unit 690 to regenerate representations of the left and right signals l, r.

In the decoder 600, the bit-stream 100 is de-multiplexed into separate streams for the dominant signal m’ for the residual signal s’ and for stereo parameters. The dominant and residual signals m’, s’ are then decoded by the decoders 620, 640 respectively. Those spectral/temporal parts of the residual signal s’ which have been encoded into the bit-stream 100 are communicated in the bit-stream 100 either implicitly, namely by detecting “empty” areas in the time-frequency plane, or explicitly, namely by means of representative signal parameters decoded from the bit stream 100. The de-correlation unit 630 and the combiner unit 650 are operable to fill empty time-frequency areas in the decoded residual signal s’ effectively with a synthetic residual signal. This synthetic signal is generated by using the decoded dominant signal m’ and output from the de-correlation unit 650. For all other time-frequency areas, the residual signal s’ is applied to construct the decoded residual signal s’. For these areas, no scaling is applied in the scaling unit 660. Optionally, for these areas, it is beneficial to transmit the aforementioned angle ρ in the encoder 500 instead of IID and coherence parameter data as data rate required to convey the single angle parameter ρ is less than required to convey equivalent IID and coherence parameter data. However, transmission of the angle ρ parameter in the bit stream 100 instead of the IID and ρ parameter data renders the encoder 500 and decoder 600 non-backwards compatible with regular conventional Parametric Stereo (PS) systems which utilize such IID and coherence parameter data.

The selector units 40, 530 of the encoders 10, 500 respectively are preferably arranged to employ a perceptual model when selecting which time-frequency areas of the residual signal s need to be encoded into the bit-stream 100. By coding various time-frequency aspects of the residual signal s in the encoders 10, 500, it is possible to thereby achieve bit-rate scalable encoders and decoders. When layers in the bit-stream 100 are mutually dependent, coded data corresponding to perceptually most relevant time-frequency aspects are included in a base layer included in the layers, with perceptually less important data moved to refinement or enhancement layers included in the layers; “enhancement layer” is also referred to as being “refinement layer”. In such an arrangement, the base layer preferably comprises a bit stream corresponding to the dominant signal m, a first enhancement layer comprises a bit stream corresponding to stereo parameters such as aforementioned angles ρ, ϕ1, ϕ2, and a second enhancement layer comprises a bit stream corresponding to the residual signal s.

Such an arrangement of layers in the bit-data 100 allows for the second enhancement layer conveying the residual signal s to be optionally lost or discarded; moreover, the decoder 600 illustrated in FIG. 10 is capable of combining decoded remaining layers with a synthetic residual signal as described in the foregoing to regenerate a perceptually mean-
ingful residual signal for user appreciation. Furthermore, if the decoder 600 is optionally not provided with the second decoder 640, for example due to cost and/or complexity restrictions, it is still possible to decode the residual signal s albeit at reduced quality.

Further bit rate reductions in the bit stream (bs) 100 in the foreground are possible by discarding encoded angle parameters \( \varphi_1, \varphi_2 \) therein. In such a situation, the phase rotation unit 680 in the decoder 600 reconstructs the regenerated output signals \( I', r' \) using a default rotation angles of fixed value, for example zero value; such further bit rate reduction exploits a characteristic that the human auditory system is relative phase-insensitive at higher audio frequencies. As an example, the parameters \( \varphi_2 \) are transmitted in the bit stream (bs) 100 and the parameters \( \varphi_1 \) are discarded therefrom for achieving bit rate reduction.

Encoders and complementary decoders according to the invention described in the foregoing are potentially useable in a broad range of electronic apparatus and systems, for example in at least one of: Internet radio, Internet streaming, Electronic Music Distribution (EMD), solid state audio players and recorders as well as television and audio products in general.

Although a method of encoding the input signals \( (l, r) \) to generate the bit-stream 100 is described in the foregoing, and complementary methods of decoding the bit-stream 100 elucidated, it will be appreciated that the invention is capable of being adapted to encode more than two input signals. For example, the invention is capable of being adapted for providing data encoding and corresponding decoding for multi-channel audio, for example 5-channel domestic cinema systems.

In the accompanying claims, numerals and other symbols included within brackets are included to assist understanding of the claims and are not intended to limit the scope of the claims in any way.

It will be appreciated that embodiments of the invention described in the foregoing are susceptible to being modified without departing from the scope of the invention as defined by the accompanying claims.

Expressions such as "comprise", "include", "incorporate", "contain", "is" and "have" are to be construed in a non-exclusive manner when interpreting the description and its associated claims, namely construed to allow for other items or components which are not specifically defined also to be present. Reference to the singular is also to be construed to be a reference to the plural and vice versa.

The invention claimed is:

1. A method of encoding a plurality of input signals \( (l, r) \) to generate corresponding encoded data, the method comprising steps of:

(a) processing, using a first processor, the input signals \( (l, r) \) to determine first parameters \( (\varphi_2) \) describing at least one of relative phase difference and temporal difference between the signals \( (l, r) \), and applying these first parameters \( (\varphi_2) \) to process the input signals to generate corresponding intermediate signals;

(b) processing, using a second processor, the intermediate signals and/or the input signals \( (l, r) \) to determine second parameters describing rotation of the intermediate signals required to generate a dominant signal \( (m) \) and a residual signal \( (s) \), said dominant signal \( (m) \) having a magnitude or energy greater than that of the residual signal \( (s) \), and applying these second parameters to process the intermediate signals to generate the dominant \( (m) \) and residual \( (s) \) signals;

(c) quantizing, in a quantizer, the first parameters, the second parameters, and encoding at least a part of the dominant signal \( (m) \) and the residual signal \( (s) \) to generate corresponding quantized data; and

(d) multiplexing, using a multiplexer, the quantized data to generate the encoded data.

2. The method as claimed in claim 1, wherein only a part of the residual signal \( (s) \) is included in the encoded data.

3. The method as claimed in claim 2, wherein the encoded data also includes one or more parameters indicative of which parts of the residual signal are included in the encoded data.

4. The method as claimed in claim 1, wherein steps (a) and (b) are implemented in the first and second processors by complex rotation with the input signals \( (l[n], r[n]) \) represented in the frequency domain \( ([k], k[k]) \).

5. The method as claimed in claim 4, wherein steps (a) and (b) are performed independently on sub-bands of the input signals \( (l[n], r[n]) \).

6. The method as claimed in claim 5, wherein sub-bands not encoded by the method are encoded using alternative coding techniques.

7. The method as claimed in claim 1, wherein, in step (c), said method includes a step of manipulating the residual signal \( (s) \) by discarding perceptually non-relevant time-frequency information present in the residual signal \( (s) \), said manipulated residual signal \( (s) \) contributing to the encoded data and said non-relevant information corresponding to selected portions of a spectro-temporal representation of the input signals \( (l, r) \).

8. The method as claimed in claim 1, wherein the second parameters in step (b) are derived by minimizing the magnitude or energy of the residual signal \( (s) \).

9. The method as claimed in claim 1, wherein the second parameters are represented by way of inter-channel intensity difference parameters and coherence parameters (IID, p).

10. The method as claimed in claim 1, wherein the second parameters are represented by way of a rotation angle \( \alpha \) and an energy ratio of the dominant \( (m) \) and residual \( (s) \) signals.

11. The method as claimed in claim 1, wherein, in steps (c) and (d), the encoded data is arranged in layers of significance, said layers including a base layer conveying the dominant signal \( (m) \), a first enhancement layer including first and/or second parameters corresponding to stereo inputting parameters, a second enhancement layer conveying a representation of the residual signal \( (s) \).

12. The method as claimed in claim 11, wherein the second enhancement layer is further subdivided into a first sub-layer for conveying most relevant time-frequency information of the residual signal \( (s) \) and a second sub-layer for conveying less relevant time-frequency information of the residual signal \( (s) \).

13. A computer-readable medium having a program recorded thereon, said program causing computing hardware to execute the method as claimed in claim 1.

14. An encoder for encoding a plurality of input signals \( (l, r) \) to generate corresponding encoded data, the encoder comprising:

(a) first processing means for processing the input signals \( (l, r) \) to determine first parameters \( (\varphi_2) \) describing at least one of relative phase difference and temporal difference between the input signals \( (l, r) \), the first processing means being operable to apply these first parameters \( (\varphi_2) \) to process the input signals to generate corresponding intermediate signals;

(b) second processing means for processing the intermediate signals and/or the input signals \( (l, r) \) to determine second parameters describing rotation of the intermediate
ate signals required to generate a dominant signal (m) and a residual signal (s), said dominant signal (m) having a magnitude or energy greater than that of the residual signal (s), the second processing means being operable to apply these second parameters to process the intermediate signals to generate the dominant (m) and residual (s) signals;

(c) quantizing means for quantizing the first parameters (φ2), the second parameters (α; IID, ρ), and at least part of the dominant signal (m) and the residual signal (s) to generate corresponding quantized data; and

(d) multiplexing means for multiplexing the quantized data to generate the encoded data.

15. The encoder as claimed in claim 14, including processing means for manipulating the residual signal (s) by discarding perceptually non-relevant time-frequency information present in the residual signal (s), said manipulated residual signal (s) contributing to the encoded data and said perceptually non-relevant information corresponding to selected portions of a spectro-temporal representation of the input signals.

16. The encoder as claimed in claim 14, wherein the residual signal (s) is manipulated, encoded and multiplexed into the encoded data.

17. A method of decoding encoded data to regenerate corresponding representations of a plurality of input signals (F', r'), said input signals (I, r) having previously been encoded to generate said encoded data, the method comprising steps of:

(a) de-multiplexing, using a demultiplexer, the encoded data (100) to generate corresponding quantized data;

(b) processing, using a first processor, the quantized data to generate corresponding first parameters (φ2), second parameters (α; IID, ρ), and at least a dominant signal (m) and a residual signal (s), said dominant signal (m) having a magnitude or energy greater than that of the residual signal (s);

(c) rotating, using a second processor, the dominant (m) and residual (s) signals by applying the second parameters (α; IID, ρ) to generate corresponding intermediate signals; and

(d) processing, using a third processor, the intermediate signals by applying the first parameters (φ2) to regenerate representations of said input signals (I, r), the first parameters (φ2) describing at least one of relative phase difference and temporal difference between the signals (I, r).

18. The method as claimed in claim 17, including in step (b) a further step of appropriately supplementing missing time-frequency information of the residual signal (s) with a synthetic residual signal derived from the dominant signal (m).

19. The method as claimed in claim 17, wherein the encoded data includes parameters indicative of which parts of the residual signal (s) are encoded into the encoded data.

20. The method as claimed in claim 17, wherein the decoder decodes parts of the encoded signal requiring supplementation by detecting empty areas of the encoded signal when represented in a time/frequency plane.

21. The method as claimed in claim 17, wherein the decoder decodes parts of the encoded signal requiring replacement or supplementation by detecting data parameters indicative of empty areas.

22. A computer-readable medium having a program recorded thereon, said program causing computing hardware to execute the method as claimed in claim 17.

23. A decoder for decoding encoded data to regenerate corresponding representations of a plurality of input signals (F', r'), said input signals (I, r) having been previously encoded to generate the encoded data, the decoder comprising:

(a) de-multiplexing means for de-multiplexing the encoded data to generate corresponding quantized data;

(b) first processing means for processing the quantized data to generate corresponding first parameters (φ2), second parameters (α; IID, ρ), and at least a dominant signal (m) and a residual signal (s), said dominant signal (m) having a magnitude or energy greater than that of the residual signal (s);

(c) second processing means for rotating the dominant (m) and residual (s) signals by applying the second parameters (α; IID, ρ) to generate corresponding intermediate signals; and

(d) third processing means for processing the intermediate signals by applying the first parameters (φ2) to generate corresponding input signals (I, r), the first parameters (φ2) describing at least one of relative phase difference and temporal difference between the signals (I, r).

24. The decoder as claimed in claim 23, wherein the second processing means is operable to generate a supplementary synthetic residual signal derived from the decoded dominant signal (m) for providing information missing from the decoded residual signal (s).

25. The decoder as claimed in claim 24, wherein the first processing means is operable to determine which parts of the residual signal (s) have been decoded for synthesizing missing non-decoded parts of the residual signal for generating substantially the entire residual signal (s).